



## 基于机器学习的广彩瓷图案生成系统

陈思齐<sup>1</sup>, 崔慧<sup>1</sup>, 杜茗晗<sup>1</sup>, 付铁明<sup>2</sup>, 孙晓红<sup>2</sup>, 纪毅<sup>†‡2</sup>, 杜本麟<sup>1</sup>

<sup>1</sup> 乐卓博大学乐卓博大学计算机科学与信息技术系, 澳大利亚墨尔本, 3086

<sup>2</sup> 广东工业大学艺术与设计的学院, 中国广州, 510006

<sup>†</sup>Email: jiyi001@hotmail.com

投稿日期: 2019-08-06; 录用日期: 2019-11-18; Crosschecked: 2019-12-12

**摘要:** 由于广彩瓷的设计元素和图案种类繁多、复杂多样, 因此准确识别现代及传统瓷器风格是广彩瓷传承工作中的巨大挑战。本文提出一种基于广彩瓷风格识别和图像合成模块的图案生成系统。在识别模块中, 通过实施主成分分析和提出的判别冗余量化策略, 对特征的重要性进行分析和排序, 然后分别训练两组神经网络, 将最优设计特征与转换后的主成分特征关联起来, 最后利用整体神经网络的逻辑回归方法对未知广彩瓷进行预测。而合成模块是基于条件生成对抗网络(cGAN)开发的, 要求用户提供自己设计的创意掩码或抽象的瓷元素图像, 以生成新的广彩瓷风格合成图像。在系统开发的过程中, 使用了603幅广彩瓷图像对分类模型进行了测试, 本方法提出的模型在精确度、召回率、接受者操作特性曲线(ROC)的曲线下面积(AUC)和混淆矩阵等方面均优于其他方法。本文研究了由用户设计的各种元素合成图像的案例, 结果表明, 该系统有助于提高学习者对广彩瓷的欣赏和艺术创作能力。

**关键词:** 广彩瓷; 分类; 生成对抗网络; 艺术创作

本文译自 Chen SS, Cui H, Du MH, et al., 2019. Cantonese porcelain classification and image synthesis by ensemble learning and generative adversarial network. *Front Inform Technol Electron Eng*, 20(12):1632-1643.

<https://doi.org/10.1631/FITEE.1900399>

中图分类号: TP751

### 1 介绍

文化遗产代表着祖先的生活方式, 为了保护这种珍贵且无法替代的资源, 人们和政府付出了巨大的努力和高昂的费用(Lowenthal, 2005)。文化遗产可分为非物质文化遗产和物质文化遗产两大类(Smith and Akagawa, 2008)。非物质文化遗产代表着保存下来代代相传的知识和技能, 包括社会实践、表演艺术、传统工艺等(Kurin, 2004)。瓷器是中国非物质文化遗产中价值最高的工艺形态之一(Ji et al., 2019)。与其他类型的中国瓷器不

同, 广彩瓷是低温上釉, 并由金编织装饰而成(中国非物质文化遗产网, 2008)。然而, 广彩瓷技术的传承极具挑战性, 年轻一代对这种工艺越来越不感兴趣。因此, 为了避免其消失, 推广与宣传广彩瓷文化的工作极为必要。

在助力广彩瓷传承工作的过程中, 瓷器种类和图案的鉴别是必不可少的重要环节。然而瓷器风格的鉴定是一个具有挑战性的问题, 这是因为广彩瓷器的器物、元素和纹样种类繁多, 其识别需要特定的领域知识。即使是经验丰富的手工艺人, 同行之间都有可能会有不同的判断。例如, 一个花瓶可能有现代图案的斗方元素、边角间隔元素、锦地元素和图案元素, 但该花瓶的风格被归类为传统。即使在边角间隔元素上有相同的图案, 花瓶和盘子也可分为不同的风格。

<sup>‡</sup> 通讯作者

\* 本文得到以下项目资助: 广州市科技计划项目(No. 2018GZMZYB17)

ORCID: 陈思齐, <http://orcid.org/0000-0001-6019-7034>

随着计算机科学技术的发展, 计算模型在分类和合成图像生成方面引起了人们极大的兴趣。计算模型在对特征之间复杂的非线性关系建模中是有效的, 它为视觉艺术领域创造性作品的识别探索提供了新的途径。特征提取是艺术绘画分类的常用方法之一。Zujovic 等人 (2009) 在特征提取领域提出一种简单而有效的方法, 用于对不同的绘画风格进行分类。神经网络在计算艺术分类中也发挥着重要作用。例如, Lecoutre 等人 (2017) 评估了各种神经网络在艺术风格识别中的作用, 提出一种残差神经网络模型用于艺术风格分类。与西方绘画相比, 中国传统绘画表现出更丰富而多样的艺术风格。Meng 等人 (2018) 在 3 种风格的中国画中训练了几种经典的卷积神经网络, 通过编辑特定的网络, 达到了很高的预测精度。Jiang 等人 (2006) 提出另一种基于 SVM 的中国画分类器, 该分类器利用低层特征来实现高层分类任务。Bao 等人 (2010) 提出一种从中国画中提取脚本的算法来揭示一幅画的意境。注意, 此算法还可用于标识和身份验证。然而, 由于现代和传统瓷器之间存在着巨大的等级差异和沉浸式设计元素, 广彩瓷的自动化分类仍然是一项艰巨的任务。

通过图像到图像的转换生成合成图像可以由传统的特定目的机器 (Efron and Freeman, 2001; Hertzmann et al., 2001; Buades et al., 2005) 和卷积神经网络 (CNN) 实现。虽然学习过程是自动化的, 但仍需花费很大精力完成人工设计各种高效的损失函数并优化, 帮助 CNN 确定什么必须最小化 (Isola et al., 2017)。Goodfellow 等人 (2014) 提出的生成式对抗网络 (GAN) 可以自动学习无需手动调整的损失函数, 因此引起了研究和行业社区的兴趣。例如, 基于 GAN 的模型被用于着色 (Iizuka et al., 2016; Larsson et al., 2016; Zhang et al., 2016) 和图像改款 (Zhu et al., 2017)。在医学研究和应用方面, Emami 等人 (2018) 针对磁共振图像, 提出一种以 ResNet 为发生器, 常规 CNN 为鉴别器的 GAN 模型来生成合成 CT 图像。Dirvanauskas 等人 (2019) 提出的 HEMIGEN 模型可以从细胞发育过程中的人类胚胎图像生成细胞的通用图像。此外, 条件 GAN (cGAN) 模型可以根据模糊和不完整的信息生成新的图像 (Li et al., 2019), 并可以从图像 (El Hattami et al., 2019)

或视频 (Chen, 2019) 中删除背景。Mao 等人 (2018) 提出的基于 GAN 的图像生成模型 SIGAN 可以生成视觉相似、语义一致的图像。素描到图像模型被应用于将素描转化为肖像、口袋妖怪 (Isola et al., 2017) 或彩色卡通图像 (Liu et al., 2018)。例如, 自动油漆工模型增加了瓦瑟斯坦损失, 以帮助监督训练克服模型崩溃, 从而实现更好的收敛。无论是素描还是彩色块素描, autopainter 都可以将素描转换成生动的卡通形象 (Liu et al., 2018)。

为协助广彩瓷信息管理, 帮助大众参与了解和欣赏文物, 本文提出一套系统, 其中包括两个模块, 用于瓷器风格识别和综合。我们的第一个贡献是提出分类模块, 用于在数据标识期间提取、表示和解释基础瓷器设计特征。这是通过使用主成分分析 (PCA) 的集成模型以及鉴别和冗余量化策略实现的。我们的另一个贡献是提出广彩瓷的再创作模块, 吸引社区交互和欣赏这一非物质文化遗产。通用引入基于 cGAN 的模块, 用户可以用他们感兴趣的对象和设计元素创建抽象图像。

## 2 方法

### 2.1 数据收集

我们从岭南展厅、广州十三行博物馆和广州荔湾博物馆 (广东工业大学) 收集了 603 幅广彩瓷的图片。每个图像都与一组设计元素和模式相关联 (图 1)。

设计元素分为七种, 包括风格、器型、主题、边角间隔、锦地、斗方和组织图案。风格是分类任务的一个指标, 其中现代风格样本 279 个, 传统风格样本 324 个。每个元素的属性和对应的观察图像数量如表 1 所示。图像的分辨率为 350×350 dpi, 图像的原始尺寸为 6000×4000 像素。

### 2.2 系统架构

图 2 给出所提出系统的框架, 其中包含用于广彩瓷器风格识别和图像重建的模块。第一个模块由特征提取和瓷器风格集成分类模型组成。第二模块以用户设计的广彩瓷创意掩码元素作为输入, 基于 cGAN 模型生成合成的广彩瓷风格图像。

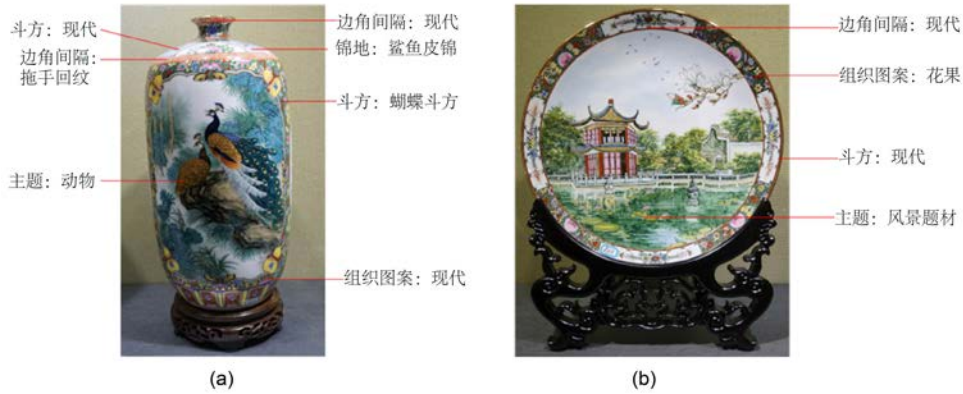


图 1 广彩瓷器不同器物图像: (a)花瓶; (b)盘子

表 1 每种元素包含的属性和观察到的广彩瓷图像的数量

元素	属性									
风格	现代 (279)	传统 (324)								
器型	碗 (20)	杯 (21)	罐 (40)	瓶 (284)	碟 (78)	盆 (6)	壶 (44)	画 (36)	其他 (74)	
主题	动物 (160)	花鸟 (8)	花 (181)	植物 (64)	当代 (11)	风景 (119)	肖像 (60)			
边角间隔	狗牙 (8)	龙爪 (84)	现代 (217)	猪鼻 (94)	如意芭蕉 (45)	鱼眼 (4)	顶工 (15)	纽绳 (8)	回纹 (4)	
锦地	水波纹锦 (3)	三线金钱锦 (8)	万字锦 (10)	四方锦 (3)	鲨鱼皮锦 (91)	三线锦 (5)	现代 (67)	胡椒锦 (14)	壁裂纹锦 (8)	
斗方	海棠斗方 (45)	现代 (95)	如意海棠斗方 (14)	蝴蝶斗方 (7)	博古斗方 (60)	指甲斗方 (19)	如意斗方 (3)	枕头斗方 (46)	抹角斗方 (4)	
组织图案	花果 (8)	现代 (218)								

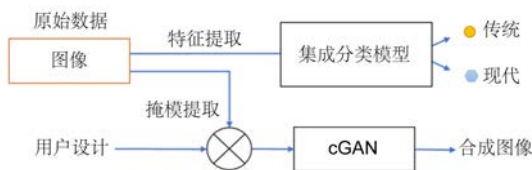


图 2 提出的广彩瓷器鉴定及合成系统的架构

### 2.2.1 瓷型分类模块

广彩瓷器风格分类预测模块如图 3 所示, 由于斗方、组织图案、边角间隔和锦地元素有多个属性, 因此需要执行 one-hot 码将这些元素转换为数值特性。总的来说, 每幅图像有  $M=35$  个特征。在该训练过程中, 为了降低特征的维数, 我们采用 PCA 和判别冗余量化策略 (特征选择 (FS)) 变换原始特征空间, 选取显著特征来表示图像样

本的特征。

用于特征表示和降维的主成分分析原理 (PCA) 执行如下: 设  $\bar{F}_{N \times M}$  为平均减去的特征, 其中  $N$  是样本的数量,  $M$  是特征的数量。通过协方差矩阵的奇异值分解获得特征向量  $\{\lambda_1, \lambda_2, \dots, \lambda_M\}$ , 并按  $\bar{F}_{N \times M}$  降序排序。相应地, 获得相应特征向量  $\{X_1, X_2, \dots, X_M\}$ 。给定要保留的  $p\%$  信息, 可以获得  $K$  个主要组件。通过将特征投影到由前  $K$  个特征向量构建的较低  $K$  维空间中, 将特征维度从  $M$  减小到  $K$ 。变换后的特征即为  $K$  个主要分量中每个分量上的样本预计权重  $W=\{w_1, w_2, \dots, w_K\}$ 。

FS 策略通过计算信息增益比 gain.ratio (Quinlan, 1986)、Relieff (Kira and Rendell, 1992)、 $\chi^2$  (Cochran, 1954) 和基于快速相关的滤波器 (FCBF) (Yu and Liu, 2003) 来对特征的重要性

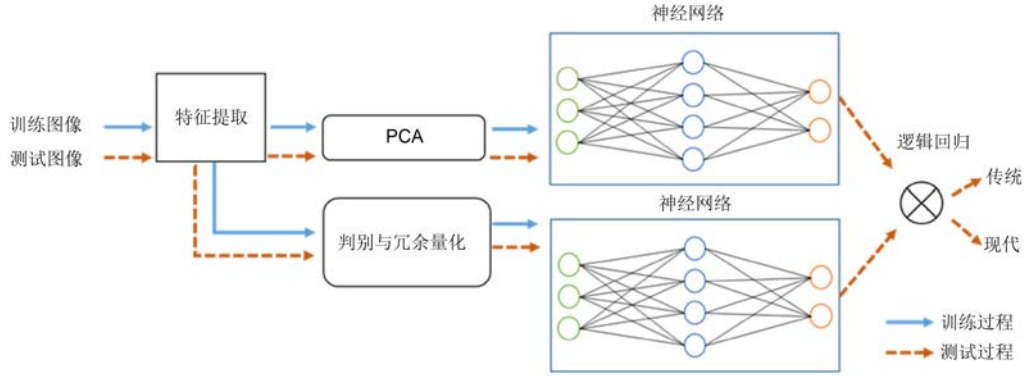


图3 广彩瓷器风格分类预测

进行排序。通过融合这 4 个指标的排名结果获得最终排名。信息增益比是信息增益和内在值之间的比率，用于在选择特征时考虑所有样本的数量和大小，从而减少对多值特征的偏差。信息增益比定义如下：

$$\text{gain.ratio}(F, f_{nm}) = \text{IG}/\text{IV}, \quad (1)$$

其中  $n=1, 2, \dots, N$ ,  $m=1, 2, \dots, M$ ,  $F$  代表特征矩阵,  $\text{IG}$  代表第  $m$  个特征,  $\text{IV}$  为内在价值。IG 和 IV 的定义如下：

$$\text{IG}(f_{n^*}, f_{*m}) = H(f_{n^*}) - \sum_{v \in \bar{f}_{*m}} \left( \frac{|\{f_{n^*} \in F \mid f_{nm} = v\}|}{|f_{n^*}|} \cdot H(\{f_{n^*} \in F \mid f_{nm} = v\}) \right), \quad (2)$$

$$\text{IV}(f_{n^*}, f_{*m}) = - \sum_{v \in \bar{f}_{*m}} \left\{ \frac{|\{f_{n^*} \in F \mid f_{nm} = v\}|}{|f_{n^*}|} \cdot \log_2 \left( \frac{|\{f_{n^*} \in F \mid f_{nm} = v\}|}{|f_{n^*}|} \right) \right\}, \quad (3)$$

其中  $f_{n^*}$  表示第  $n$  个样本的所有特征,  $f_{*m}$  表示  $F$  中的第  $m$  个特征的向量,  $H$  表示熵值。

ReliefF 算法能够区分彼此接近的数据样本, 定义为

$$d(f_{nm}, f_{ym}) < d(f_{nm}, f_{zm}), \quad (4)$$

其中  $y$  是同一类中第  $n$  个随机样本的最近邻,  $z$  是

另一类样本。如满足方程(4), 则  $f_m$  的特征权重增加, 反之亦然,  $d$  表示这两个样本之间的欧氏距离。

$\chi^2$  测试用于确定预期特征与一个或多个类中观察到的特征之间是否存在显著差异。在零假设正确的情况下, 即  $r \rightarrow \infty$ ,  $\chi^2$  分布的限制分布被定义为

$$\chi^2 = \sum_{i=1}^c \frac{(f_i - q_i)^2}{q_i} = \sum_{i=1}^c \frac{f_i^2}{q_i} - r, \quad (5)$$

其中  $C$  是相互排斥的类, 具有各自观察到的特征  $f_i$  ( $i=1, 2, \dots, C$ ),  $r$  为从分类为  $c$  个互斥类的总体的随机样本中的观察,  $p_i$  为观察值属于第  $i$  类的概率。因此, 对于所有  $i$ , 我们都有期望数  $q_i = r p_i$ , 其中  $p_i$  和  $q_i$  定义如下：

$$\sum_{i=1}^c p_i = 1, \quad (6)$$

$$\sum_{i=1}^c q_i = r \sum_{i=1}^c p_i = \sum_{i=1}^c f_i. \quad (7)$$

特征选择算法 (FCBF) 是一种多变量特征选择方法, 它从一组完整的特征开始, 使用对称不确定性 SU。SU 是一种规范化的信息理论度量, 它利用熵  $H$  和条件熵值, 通过序列搜索策略, 利用逆向选择技术计算特征相依性, 并找到最佳子集。SU=1 表示一个特征的值可以由另一个特征完全预测, 而 SU=0 表示两个特征完全独立。SU 被定义为

$$\text{SU}(f_{nm}, f_{im}) = 2 \frac{H(f_{nm}) - H(f_{nm} \mid f_{im})}{H(f_{nm}) + H(f_{im})}, \quad (8)$$

其中  $f_{nm}$  和  $f_{im}$  是随机样本的特征,  $H(f_{nm}|f_{im})$  是条件熵,  $H(f_{nm})$  是  $f_{nm}$  的熵.  $H(f_{nm}|f_{im})$  和  $H(f_{nm})$  定义如下:

$$H(f_{nm} | f_{im}) = -\sum P(f_i) \sum P(f_n | f_i) \log_2 P(f_n f_i). \quad (9)$$

$$H(f_{nm}) = -\sum P(f_n) \log_2 P(f_n), \quad (10)$$

其中  $P(f_n)$  是  $f_n$  的概率.

通过融合信息增益比率、ReliefF、FCBF 和  $\chi^2$  的归一化排名分数, 可以获得每个特征的重要性和判别力.

在训练过程中, 考虑到  $W$  和排名靠前的特征, 分别对两个神经网络进行训练. 为了更好地关联这两个模型的分类结果, 采用一个集合模型对一个未知的瓷器样品进行预测, 该模型采用逻辑回归方法对结果进行整合.

### 2.2.2 广彩瓷器再创作模块

第二模块是基于 cGAN (Isola et al., 2017) 开发的广彩瓷图像再创作模块. 在这项工作中, 研究团队提取了广彩的构成元素, 如图 4 所示, 总计 196 幅图像和 23 种不同类型的物体, 包括鹤、果树、构图、船、建筑、石头、竹子、鸟、鸡、

狗、猫、蝴蝶、渔夫等.

该训练过程如图 5 所示. 通过给定图像  $I$  和相应的掩码  $S$ , 鉴别器  $D$  学习如何鉴别由生成器  $G$  和输入  $\{I, S\}$  创建的伪图像, 而  $G$  旨在生成伪图像以欺骗鉴别器  $D$ . 与传统 GANs 不同的是, 在 cGAN 中  $G$  和  $D$  都观察输入掩码. cGAN 的目标定义为:

$$G^* = \arg \min_G \max_D L_{\text{cGAN}}(G, D) + \lambda L_{L_1}(G), \quad (11)$$

其中  $\min_G$  和  $\max_D$  分别表示生成器  $G$  试图最小化目标函数, 而鉴别器  $D$  试图最大化目标函数,  $L_{L_1}$  是由距离  $L_1$  计算的损耗.  $\lambda$  是用于调整权重  $L_{L_1}$  的参数, 表示  $G$  欺骗  $D$  的形式与传统 GAN 无异, 同时其与  $L_1$  的距离需尽可能接近原始值.  $L_{\text{cGAN}}$  是 cGAN 损失, 定义如下:

$$L_{\text{cGAN}}(G, D) = E_{I,O}[\log D(I, O)] + E_{I,U}\{\log(1 - D[I, G(I, U)])\}, \quad (12)$$

其中  $E$  是概率论中的期望值,  $O$  是输出图像,  $U$  是随机噪声矢量. 当  $G^*$  收敛或达到预定数量的时期时, 可以获得训练的收敛或达到预定数量时,



图 4 广彩瓷图案及附有标签物件的构成元素案例

树、房子、桥、船、鸟、渔夫和竹子分别以深绿色、浅蓝色、棕色、蓝色、橙色、粉色和亮绿色表示

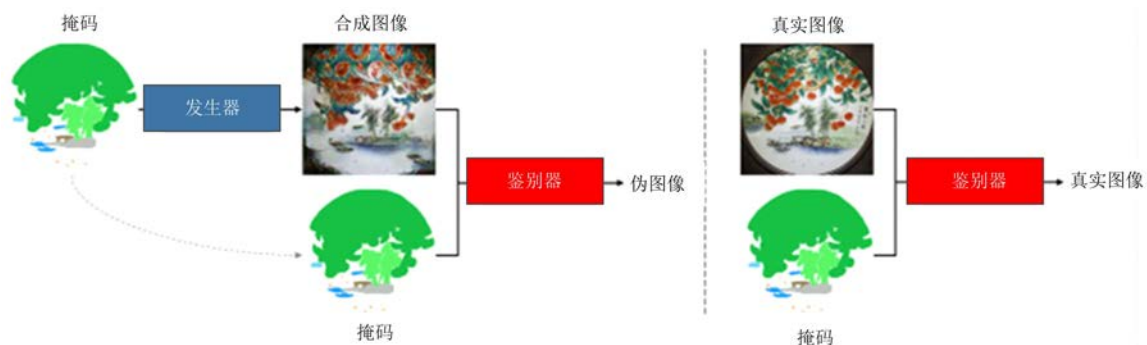


图 5 基于 cGAN 的元素-图像生成模型的训练过程

可获得训练的 cGAN 模型。

给定训练 cGAN 模型后，可以采用用户设计的创意掩码或抽象图像作为输入来获得合成图像。

### 2.3 方法比较

评估模型在广彩瓷识别中的作用，本文将 Stack-PCA-FS-NN 与 PCA-NN、FS-NN、Stack-NN 这三个模型相比较。其中 PCA-NN 是只有 PCA 的神经网络，FS-NN 是只有 FS 的神经网络，Stack-NN 是没有进行 PCA 或 FS 降维的模型。

### 2.4 评估方法和参数设置

#### 2.4.1 评价矩阵

所有分类模型均通过 10 折交叉验证进行评估。采用测试数据集的平均测试结果作为性能评判的指标。预测性能通过接受者操作特性曲线 (ROC) 的曲线下面积 (AUC)、精确度 (precision)、召回率 (recall) 和混淆矩阵来进行评估。精确度表示正确预测的阳性样本数与所有预测为阳性的样本数之比，定义如下：

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (14)$$

其中 FN 为假阴性 (错误识别的阳性样本数)。  $F_1 = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$ ，影响了精度和召回率的计算。

ROC 空间由假阳性率 (FPR) 和真阳性率 (TPR) 分别定义为  $x$  和  $y$  轴，描述了 TP 和 FP 之间的相对权衡关系。TPR 和 FPR 的定义如下：

$$\text{TPR} = \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (15)$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}, \quad (16)$$

其中 TN 为真阴性 (正确预测的阴性样本数)。ROC 的 AUC 越高，模型的分类性能越好。

#### 2.4.2 参数设置

在第 1 个分类模块中，使用 ReLu 激活神经网络，其中神经元个数设为 100，最大迭代次数为 200。为评估集成结构、PCA 和特征排序策略对分

类性能贡献，在模型简化测试中参与比较的所有方法的神经网络设置均一致。因此，在系统的原始设计中，我们没有对网络进行最佳参数的自动选择。注意，一个自动化的参数选择机制 (Polap et al., 2018) 可用于具有复杂设计元素和特性的大型数据集。

在 PCA 中将  $K$  设为 10，在特征排序策略中选取排名前 12 位的特征。参数设置的影响将在第 3 节中讨论。

在第 2 个基于 cGAN 的模块中，参数设置与 Isola 等人的 pix2pix code (Isola et al., 2017) 相同，其中 epoch 设置为 200， $\lambda$  设为 100。

## 3 结果和讨论

### 3.1 分类结果

所提出的 Stack-PCA-FS-NN 模型及所有其他模型分类结果如表 2 所示，AUC ROC 曲线如图 6 所示。总体而言，集成模型在 AUC、精确度和召回率方面均优于其他模型，PCA 的神经网络提高了 FS 神经网络的 AUC，并且也提高了精确度和召回率。基于 FS 和 PCA 的特征约简和变换的集成模型性能优于 Stack-NN。其中，ROC 曲线下面积 (AUC) 为 0.944，精确度 (Precision) 为 0.872，召回率 (recall) 为 0.872。

表 2 所提出模型和其他有/无 PCA、判别特征选择 (FS) 和集成策略方法的分类性能

模型	ROC 曲线	精确度	召回率	$F_1$
	下面积	(mean±std)	(mean±std)	
FS-NN	0.920	0.832±0.006	0.828±0.007	0.830
PCA-NN	0.936	0.840±0.010	0.833±0.007	0.837
Stack-NN	0.906	0.851±0.010	0.839±0.008	0.845
Stack-PCA-FS-NN	0.944	0.87±0.005	0.872±0.006	0.872

图 7 给出混淆矩阵方法的比较结果。研究表明，在对传统分类样本进行预测时，PCA 变换后的特征神经网络比 FS 排序前的特征神经网络表现更好，而对于现代分类样本，NN-FS 比 NN-PCA 分类效果更好。在不降低特征维数情况下，集成模型在两个类上取得较好性能。与传统分类方法相比，该方法显著提高了分类性能。

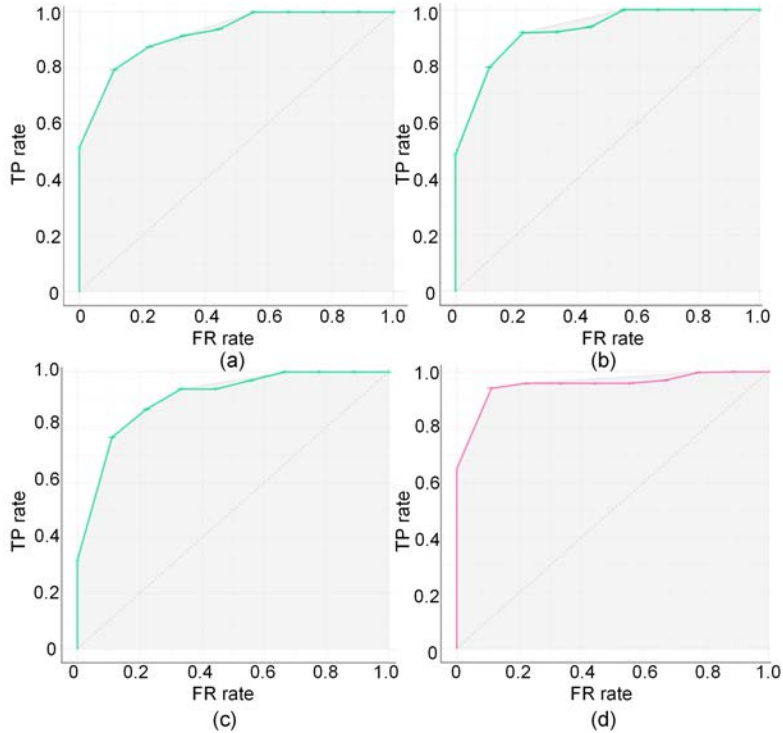


图 6 由 FS-NN (a), PCA-NN (b), Stack-NN (c) 和 Stack-PCA-FS-NN (d) 4 种方法得到的 AUC ROC 结果曲线

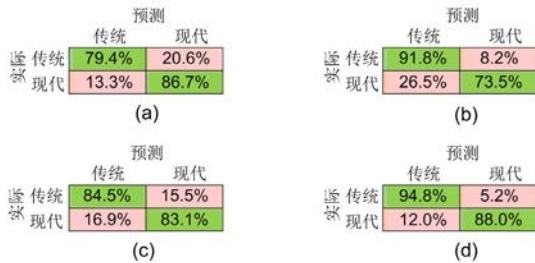


图 7 FS-NN (a), PCA-NN (b), Stack-NN (c) 和 Stack-PCA-FS-NN (d) 的混淆矩阵比较

通过选取不同的  $K$  值 (从 7 到 20), 研究主成分个数对主成分分析结果的影响。主成分分析所涵盖的 ROC 曲线 AUC、recall 以及对应的方差如图 8 所示。当主成分数量为 10 时, AUC 和召回率 recall 均达到第一个峰值。因此, 此项目中, 当所覆盖的方差为 54% 时,  $K$  的个数设为 10。

研究项目利用树形结构研究了传统类和现代类中元素、模式和特征的原始分布。如图 9 所示, 在 603 例中, 324 例属于传统类。在碗、画或壶中, 92% 的瓷器是现代的。对于杯、罐、盘、盆、瓶等器物, 62.8% 为传统器物。在这些传统瓷器中, 画中出现现代图案的, 82.5% 是传统瓷器, 其余是现

代瓷器。如果广彩图案没有很明显的现代特征, 则很难判断广彩瓷是否属于现代分类。在现代的样品中, 86.8% 的杯子、罐子、盘子或盆是传统器皿。因此, 对特征与目标之间的关系进行建模是一大挑战。对特征的识别和冗余度进行定量评价, 有助于发现显著特征。

图 10 给出所提出的判别和冗余量化策略的前 35 个特征。可以看出, 信息量最大的特征包括组织图案元素中的现代模式、边角间隔元素中的拖拽模式、斗方元素中不同模式的数量、锦地中的鲨鱼皮锦等。在排名前 12 位的特性中, 有 4 个是不同元素中不同模式的数量, 3 个是边角间隔元素中的模式。由图 7 中混淆矩阵评价的实验结果可以看出, PCA 变换后的特征有助于对传统类进行分类, 而提出的特征排序策略对现代类数据的分类效果更好。本文还研究了排名靠前的特征对分类的影响。当使用 7 到 19 个顶级特性时, AUC 和召回率结果如图 8 所示。如图 11 所示, 当选择的显著特征为 10 个时, AUC 和召回率都达到局部最小值。当使用排名前 12 位的特性时, 性能得到改善, 并达到局部最大值。当显著特征的数量大于 14 时, AUC 和召回率的分类结果仍有轻微变化。

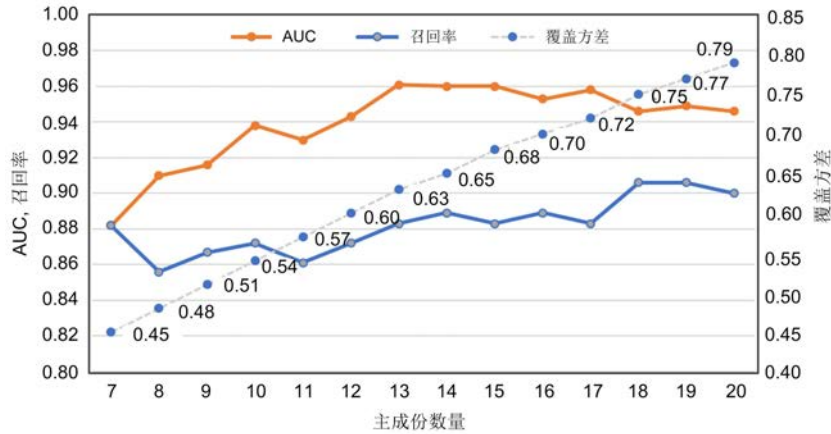


图 8 所提出的分类模型中主成分个数对主成分分析结果的影响

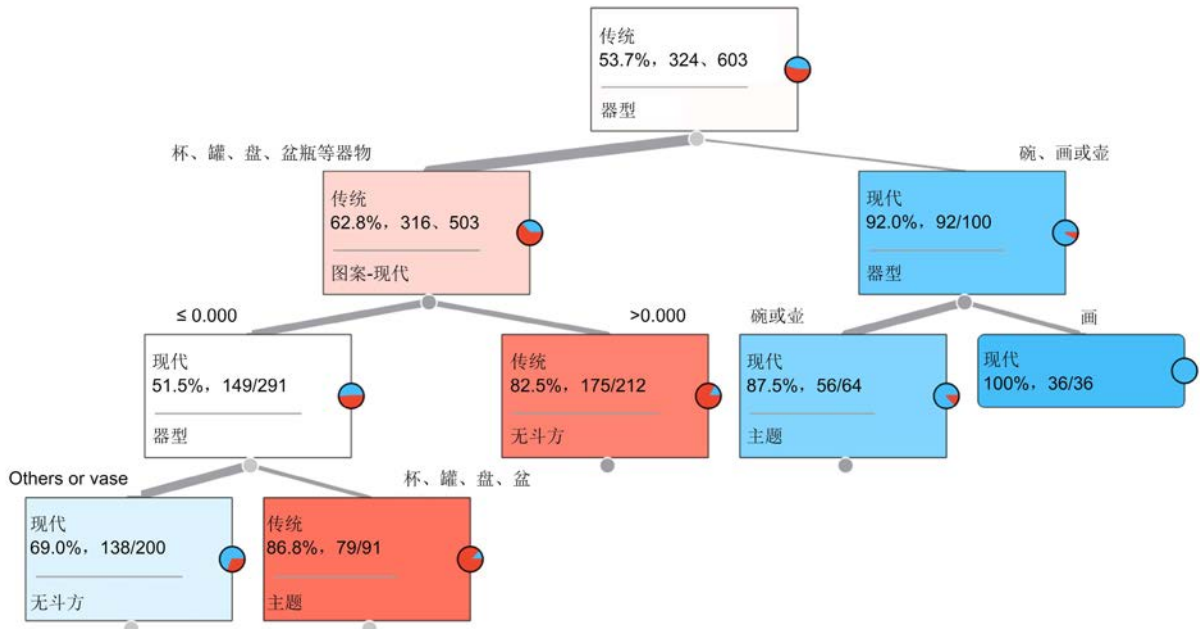


图 9 包含特征重要性的树状结构决策模式（传统类为红色，现代类为蓝色）

### 3.2 合成结果

训练过程中损失随时间的变化如图 12 所示，仅用 1 台 CPU 进行 GAN 模型培训，共花费 22 小时。

图 13 显示了合成图像的 3 个例子和相应的用户输入。给出一个简单的抽象形象与果树绘画，合成图像生成如图 13a 所示。通过添加鹤的图案和构图框架，生成添加鹤的合成图像，如图 13b 所示。最后，如果用户提供了一个更为复杂的混合绘图，其中包括了果树、船、竹子、石头、鸟类和桥梁，可以得到一个风景合成图，如图 13c

所示。如图 13 所示，即使采用简单的抽象绘画的图像，所提出系统依然可以生成广彩瓷风格的图像，并基于在训练数据集中对真实广彩瓷图像的学习成果来填充背景。

在目前的系统中，有 23 种不同类型的图纸。虽然合成的图像与真实的图像有很大的相似性，但是研究项目未来的工作包括增加训练图像的数量和标记对象，以使用户绘制更加多样化的抽象图像。

在当前系统中有 23 种不同类型对象，如鹤、果树、框架、船、建筑、石头、竹子、鸟、鸡、狗、猫、蝴蝶和渔夫。用户可以灵活选择一个或



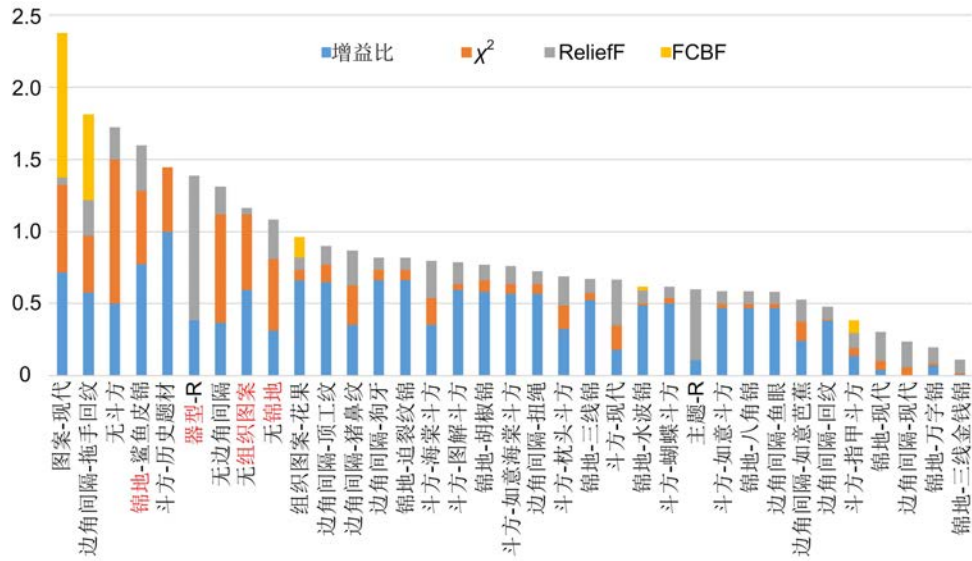


图 10 判别和冗余量化策略中信息较大特征排序  
在排名前 35 位的特性中，7 个是边角间隔特征，8 个是锦地特征

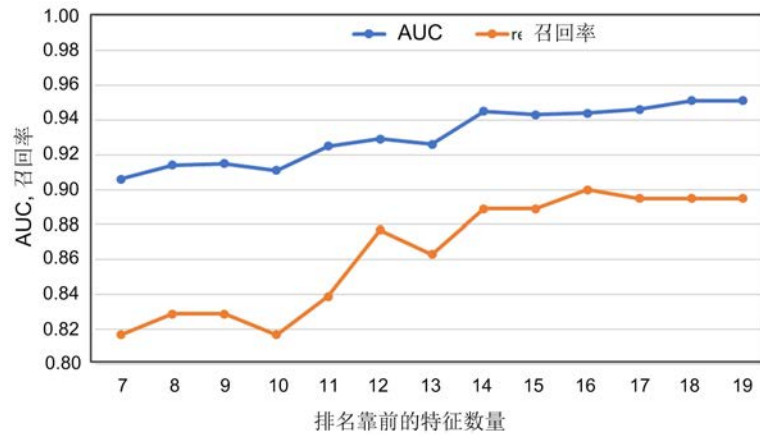


图 11 排名靠前特征数量对分类结果的影响

多个所需的对象掩模进行绘图，但仅限于系统中已有的类型。因此，通过 23 个不同物体，7 个设计元素和 35 个特征，证明了所提出的广彩瓷风格图像重建系统的有效性。造成合成图像模糊的原因之一是相对于自然图像，广彩瓷的对象比较有限。未来，我们计划从更大的公共数据集（如 COCO (Lin et al., 2014)）和更广泛的对象中实现转移学习。此外，我们计划使用自动化参数搜索方法，如多线程学习控制机制 (Polap et al., 2018)，用于引入大型数据集以及更复杂的设计元素和特性。

#### 4 结论

本文提出一套可解释广彩瓷风格的分类体系。一项模型简化测试研究的结果表明，在分类性能方面，综合 PCA 神经网络和判别特征的方法优于其他方法。该系统提供了一个模块，从用户设计的广彩瓷元素抽象图像中合成广彩瓷风格图像。个案研究以 23 件物品为样本分析了广彩瓷的合成图像，结果显示，所提出的系统在引导市民欣赏和保护这些非物质文化遗产方面有较大潜力。

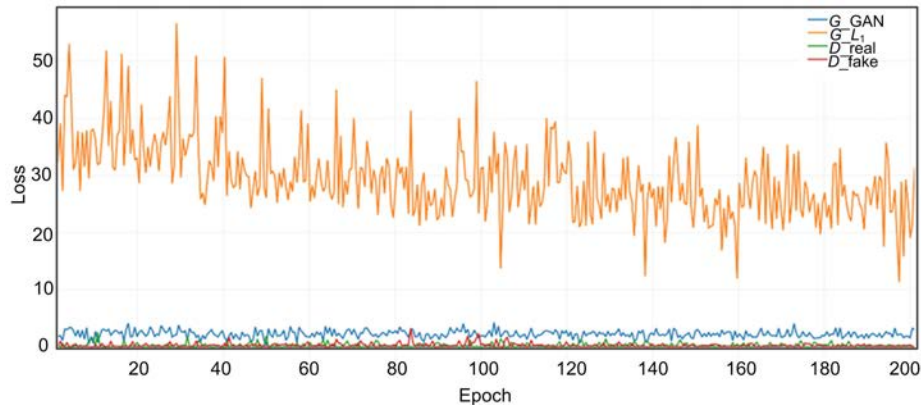


图 12 cGAN 模型训练过程中随时间变化的损失 (G 表示产生器, D 表示鉴别器)

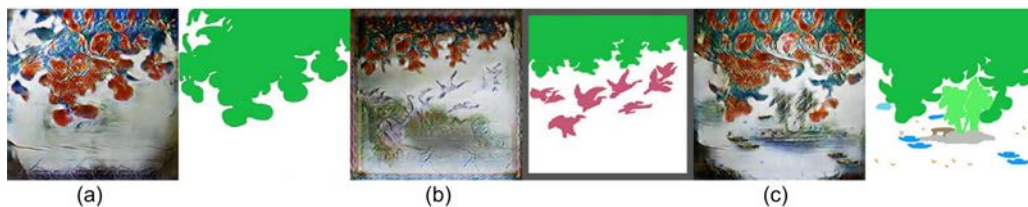


图 13 当给出简单图像 (a) 和混合图像 (b)、(c) 时, 所提出系统生成的广彩瓷风格图像的合成实例

### 遵守伦理准则声明

作者声明发表这篇论文没有利益冲突。

### 参考文献

- Bao H, Liang Y, Liu HZ, et al., 2010. A novel algorithm for extraction of the scripts part in traditional Chinese painting images. Proc 2<sup>nd</sup> Int Conf on Software Technology and Engineering, p.V2-26-V2-30. <https://doi.org/10.1109/ICSTE.2010.5608756>
- Buades A, Coll B, Morel JM, 2005. A non-local algorithm for image denoising. IEEE Computer Society Conf on Computer Vision and Pattern Recognition, p.60-65. <https://doi.org/10.1109/CVPR.2005.38>
- Chen KH, 2019. Image Operations with cGAN. <http://www.k4ai.com/imageops/index.html>
- China Intangible Cultural Heritage Network, 2008. Cantonese Porcelain Inheritance Project. [http://www.ihchina.cn/project\\_details/14453/](http://www.ihchina.cn/project_details/14453/) [Accessed on July 16, 2019] (in Chinese).
- Cochran WG, 1954. Some methods for strengthening the common  $\chi^2$  tests. *Int Biom Soc*, 10(4):417-451. <https://doi.org/10.2307/3001616>
- Dirvanauskas D, Maskeliūnas R, Raudonis V, et al., 2019. HEMIGEN: human embryo image generator based on generative adversarial networks. *Sensors*, 19(16):3578. <https://doi.org/10.3390/s19163578>
- Efros AA, Freeman WT, 2001. Image quilting for texture synthesis and transfer. Proc 28<sup>th</sup> Annual Conf on Computer Graphics and Interactive Techniques, p.341-346. <https://doi.org/10.1145/383259.383296>
- El Hattami A, Pierre-Doray É, Barsalou Y, 2019. Background removal using U-net, GAN and image matting. <https://github.com/eti-p-doray/unet-gan-matting> [Accessed on July 14, 2019].
- Emami H, Dong M, Nejad-Davarani SP, et al., 2018. Generating synthetic CTs from magnetic resonance images using generative adversarial networks. *Med Phys*, 45(8): 3627-3636. <https://doi.org/10.1002/mp.13047>
- Goodfellow IJ, Pouget-Abadie J, Mirza M, et al., 2014. Generative adversarial nets. Proc 27<sup>th</sup> Int Conf on Neural Information Processing Systems, p.2672-2680.
- Hertzmann A, Jacobs CE, Oliver N, et al., 2001. Image analogies. Proc 28<sup>th</sup> Annual Conf on Computer Graphics and Interactive Techniques, p.327-340. <https://doi.org/10.1145/383259.383295>
- Iizuka S, Simo-Serra E, Ishikawa H, 2016. Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. *ACM Trans Graph*, 35(4), Article 110. <https://doi.org/10.1145/2897824.2925974>
- Isola P, Zhu JY, Zhou TH, et al., 2017. Image-to-image translation with conditional adversarial networks. IEEE Conf on Computer Vision and Pattern Recognition, p.1125-1134. <https://doi.org/10.1109/CVPR.2017.632>
- Ji Y, Tan P, Chen SC, et al., 2019. Kansei engineering for E-commerce Cantonese porcelain selection in China. 21<sup>st</sup> Int Conf on Human-Computer Interaction, p.463-474. [https://doi.org/10.1007/978-3-030-22646-6\\_34](https://doi.org/10.1007/978-3-030-22646-6_34)

其余文献从略